崇城大学　紀要
第44巻　平成31年3月

An Exploratory Factor Analysis of Scores on the Feelings Towards Group Work Scale in the Japanese University EFL Context

by
Larry XETHAKIS*

Abstract

The Feelings Towards Group Work questionnaire (FTGW; Cantwell & Andrews, 2002) is an instrument that has been used in a wide range of contexts and populations to examine learners' dispositions towards group learning. Using data from 307 Japanese university students, this study reports on an exploratory factor analysis (EFA) conducted on the FTGW. The purpose of the EFA solutions executed was to gain post hoc insight into the structure of scores generated by a version of the FTGW adapted for use in the Japanese context. An earlier study using confirmatory factor analysis (CFA) conducted by the author (Xethakis, 2018) revealed unsatisfactory fit on the three-factor model hypothesized by the authors of the instrument. The results of the EFA (1) revealed that the dimensionality of the scores gathered in this data set points to a lack of coherence in general rather than a coherence different from Cantwell and Andrews (2002); and (2) suggests a path for this adapted version be adjusted in any future research. The results also raise the theoretical issue of whether preference for group work and preference for individual work should be treated as two separate constructs, or as opposite poles on a single construct. Preliminary evidence in this study suggests the latter.

Key Words：collaborative learning, group work, validity

1. Introduction

The literature on group-based learning approaches ascribes a wide-range of benefits to this family of pedagogical techniques. Exposure to approaches such as collaborative learning has been shown to develop learners' academic skills, build their self-esteem and improve their overall attitudes towards learning, in addition to enhancing their interactions with their peers (Johnson, Johnson & Smith, 1998; Slavin, 1996). In addition to these more general benefits, Gillies (2003) notes that peers can provide assistance by explaining new ideas and unfamiliar information, something of great importance considering the trend towards more student-centered forms of instruction. In the realm of language learning, Dornyei and Malderez (1997) report that group members, in addition to their role as part of a pool of language resources, can act as sources of motivation and form a base of support for others. Perhaps it is because of these kinds of benefits that the MEXT has specifically promoted the use of group-based learning approaches among other

＊ Senjor Assistant Professor, Sojo International Learning Center

innovative approaches in its guidelines for improving the quality of English education in Japan (MEXT, 2008; 2015).

While group-based learning approaches provide students with a number of benefits that are not found in more traditional, teacher-centered learning approaches, the use of such techniques in the classroom also brings with it a number of potential issues. One such issue is that teachers employing group-based learning approaches make a pair of assumptions, one affective and one cognitive, that underlie the use of group work in the classroom (Cantwell & Andrews, 2002). The first of these is the assumption that “all children are comfortable with the idea and processes” (p. 76) of group work. For some learners, particularly those who feel a sense of anxiety in social situations, working in a group, rather than providing support and motivation, may be a source of discomfort, and in more extreme cases, even of distress (Fantuzzo, Riggio, Connelly, & Dimeff, 1989; Kondo & Yang, 1994; Zhou, 2016). The second of these assumptions is that learners also possess the necessary cognitive and metacognitive attributes to function successfully in groups. One of the requirements for working in groups is that learners be able to manage the complexity associated with social interaction and the switching of roles. For example, learners may have to shift between acting as a teacher to peers and then asking for assistance, while meeting other simultaneous requirements calling for social dexterity and sensitivity. Cantwell and Andrews (2002) suggest that for students with a lesser degree of metacognitive awareness negotiating this complexity may prove difficult.

In group-based learning approaches, individual differences in personality traits and level of cognitive development among group members take on a more significant role in learning outcomes than they would in a more traditionally structured and teacher-centered classroom. For learners whose cognitive or psychological dispositions leave them unsuited or unprepared to work in groups, group-based learning approaches, rather than bringing benefits, might actually create obstacles to learning. Gorvine and Smith (2015), in a study examining American university-level learners’ attitudes towards group work and aspects of their attitudes toward learning statistics suggest that teachers should gain an understanding of how “students’ preexisting attitudes may influence their engagement with the material and final performance in the course” (p. 56). White, Lloyd, Kennedy, and Stewart (2005) express similar sentiments in stressing the importance, for both practitioners and researchers, of gaining “an understanding of how students feel towards group work” (p. 617). These observations may be even more relevant for teachers and researchers working with group-based learning approaches in L2 classrooms, considering the increased psychological and cognitive demands made of learners in such situations.

While there has been research into the role that psychological and cognitive factors play in students’ dispositions towards group work outside of Japan (e.g., Cantwell & Andrews, 2002; Fantuzzo, et al., 1989; Forrester & Taschian, 2010; Gorvine & Smith, 2015; Zhou, 2016), only a limited amount of research has been carried out in the Japanese context (e.g., Fushino, 2010; Nosaki, 2016). Moreover, the instruments used in these Japanese-context studies were developed for the particular studies themselves, and their psychometric properties have not been established on the basis of independent research, nor were they, in some cases (e.g., Nosaki, 2012, 2016), established in the original studies themselves. For this reason, there is a need for an evidence-based measure which examines students’ dispositions towards working in groups in the Japanese context. Such an instrument would allow researchers to explore this area with greater confidence in their findings, while at the same providing educators with a means to help identify students who, as a matter of individual differences, might well be handicapped when working in group-based learning situations.

One of the most relevant studies in this area of research, i.e. concerning the impact of learners' individual differences and their dispositions towards group work, was undertaken in 2002 by Cantwell and Andrews. In this study, the authors hypothesized that a number of cognitive and psychological factors, such as social anxiety, fear of negative evaluation, achievement goals, metacognitive awareness, and need for affiliation, might underlie learners' dispositions towards group learning. To investigate this hypothesis, Cantwell and Andrews first developed an instrument, using exploratory factor analysis (EFA), to measure such dispositions. The EFA supported a three-factor structure underlying the scores derived in a pilot study of the instrument. This latent structure, advocated by Cantwell and Andrews, comprised factors apparently representing preference for individual learning, preference for group learning and discomfort in group learning.

These three factors were adopted as distinct subscales of the newly-developed instrument, which the authors termed the Feelings Towards Group Work (FTGW) questionnaire (see Appendix), and the questionnaire has 18 items in total. The first of the aforementioned subscales is comprised of seven items and was labeled the Preference for Individual Learning (PIL) subscale. The items on this subscale generally focus on an individual's dislike of group work as well as feelings with respect to its ineffectiveness as a means of learning. An example item from this subscale is "*I often think the work becomes too confusing when done in a group rather than individually*." The second subscale, designated the Preference for Group Learning (PGL) subscale, is also comprised of seven items. This subscale is generally focused on the interdependent nature of group work as well as feelings of success when such a condition is realized in the group. Item 7, "*I often have a strong feeling of satisfaction when I become totally involved in a group achievement*," is an example of this focus. The third subscale was termed the Discomfort in Group Learning (DGL) subscale and is comprised of only four items. These items express feelings of unease at being in the group and with a number of the processes necessary for group work, such as sharing ideas or opinions, asking for help or, as in Item 4, gaining a clear understanding of the group task. Responses to all items are on a Likert scale, with values ranging from 1 to 5. Cronbach's alpha for each of the three subscales were reported in the original study as .78, .71 and .60, respectively.

Cantwell and Andrews then employed this newly-developed instrument together with a number of other instruments, including the Fear of Negative Evaluation (FNE) and Social Avoidance and Distress (SAD) Scales (Watson & Friend, 1969), the Metacognitive Awareness Inventory (MAI, Schraw & Dennison, 1994), the Achievement Goals Questionnaire (Archer, 1994) and the Need for Affiliation subscale from Personality Research Form E (Jackson, 1974), to examine the possible relationships between the cognitive and psychological factors measured by these instruments and respondents' dispositions towards group-based learning approaches. In the Cantwell and Andrews study, the target population comprised secondary students, with students from years seven, nine and eleven sampled as representatives of this educational range. As a result of their research, Cantwell and Andrews found positive correlations between the PIL subscale of the FTGW and both the SAD and the FNE. Respondents with high scores on these two anxiety-measuring instruments were found to have similarly high scores on the DGL subscale as well. Respondents' scores on the PGL subscale correlated negatively with these two instruments and positively with scores on the MAI, the presence of mastery or performance goals, and with a need for affiliation.

Since its inception, the FTGW has moved beyond its original population and context into university settings and non-English speaking populations. At the university level, the most recent study is that of Gorvine and Smith (2015) mentioned above. This study examined students' attitudes towards working in groups, anxieties toward learning statistics and learning outcomes using the FTGW, along with an instrument designed to measure attitudes toward studying statistics developed for the study. They found that high scores on the PGL subscale of the FTGW and lower degrees of anxiety towards learning statistics were correlated with better learning outcomes. From this they inferred that learners who felt more comfortable in group settings would be more inclined to take advantage of the learning opportunities available in groups, and thus would perform better in a course. In a 2010 study, Forrester and Taschian examined the role that five widely-studied personality traits ― agreeableness, conscientiousness, neuroticism, extraversion, and openness to experience ― had on the dispositions of students majoring in business at an American university. To this end, two of the three subscales of the FTGW (the PGL and DGL) together with the Neo Five-Factor Inventory (Neo-FFI; Costa & McCrae, 1992) were employed. In this study, a positive correlation was found between high scores on the PGL subscale and the personality trait of extroversion. The authors also found a positive correlation between the trait of neuroticism and scores on the DGL subscale, as well as a slight negative correlation between scores on this scale and extraversion. Moreover, learners who scored high in extroversion were able to function successfully in groups, whereas those who scored high in neuroticism were limited in their ability to do so. One additional study at the university level employing the FTGW is that of White, et al. in 2005. The respondents in this study were students in IT and Pharmacology programs, and the aim was to examine learners' attitudes towards working in groups and group assessment. The authors employed an adapted version of the FTGW and an initial overall preference for group work was found in students from both programs. Students in Pharmacology also reported a slight increase in their preference for group work in their post-intervention responses.

Outside of English-speaking populations, the most extensive use of the FTGW has been in Greece. In 2009, Goudas, Magotsiou and Hatzigeorgiadis examined the psychometric properties of scores on a substantially modified Greek-language version of the FTGW in a sample population of Greek sixth-grade students. As a result of confirmatory factor analysis (CFA), Goudas et al., found that the data set in their study had a three-factor structure similar to that found in Cantwell and Andrews' original study. They re-named this modified version of the instrument the G-FTGW, and it has been used in a number of additional studies (Angeli & Tsaggari, 2016; Goudas & Magotsiou, 2009; Lemonia & Dimitris, 2017) in similar populations. In addition to the Greek version, the FTGW has also been translated into Arabic. Gasaymeh, Kreishan and Al-Dhaimat (2014) utilized the PGL and DGL subscales of the FTGW to investigate preferences for group learning among students in the education and computer science departments of a Jordanian university. Respondents from both departments expressed a preference for group learning, with the students from the computer science department expressing a stronger preference for this approach.

The FTGW has also been used in the university EFL context which is also the context of the study reported in this paper. Maesin, Mansor, Shafie, and Nayan (2009) made use of an adapted version of the FTGW to examine the learning preferences of science and social science majors at a Malaysian university who were enrolled in English classes. The authors did not, however, report in their study whether the instrument was translated or used in the original English form. The respondents in this study expressed a preference for the use of group learning approaches in their English classes, with social science majors expressing a slightly higher preference for this teaching method.

Because of its use in a wide range of contexts and populations, the FTGW was thought to be a good candidate for adaption into the Japanese EFL context by the author of this paper (see Xethakis, 2018 for a discussion of the adaptation process). However, the results of a CFA on scores generated by this adapted version, in a sample of students in university-level EFL classes, showed that the three-factor model originally hypothesized by Cantwell and Andrews' had unambiguously poor-fit with the data. In further diagnostic CFAs carried out separately on each of the FTGW's subscales in the same study (Xethakis, 2018), it was also found that while the DGL subscale had meritoriously good fit with the dimensionality of the data, the two remaining subscales, the PIL and PGL, exhibited poor-fit with the data. In the current study, the results of an EFA, executed in pursuit of an a posteriori analysis of the dimensionality of the same the data as that used in Xethakis (2018), are reported. This represents a new set of analyses under the rationale of clarifying what the structure of scores generated by the Japanese adapted version might actually be, if not commensurate with Cantwell and Andrews' originally hypothesized three-factor structure.

2. Methodology

The methodology for this study is described in three sections below. First, the characteristics of the instrument are described, followed by those of the participants in the study and the process of data collection. The third section presents the procedures followed in the statistical and psychometric analysis of the data, including ascertainment of normality, the EFA, which represents the central contribution of this paper, and finally, the calculation of reliability estimates (Cronbach's alpha) for each of the subscales revealed by the EFA.

2.1 Instrument

Cantwell and Andrews' FTGW (2002) is an 18-item instrument that consists of three hypothesized subscales. These were taken by its authors to measure three distinct facets of respondents' feelings towards group work — the PIL subscale (Items 1, 5, 6, 12, 14, 16 and 18), the PGL subscale (Items 3, 7, 8, 9, 11, 13 and 15) and the DGL subscale (Items 2, 4, 10 and 17). The latter subscale (four items) is shorter than the first two subscales (seven items). The scales are based on a semantically anchored Likert scale, ranging from 1, *not true of me at all*, to 5, *very true of me* (see Appendix for the FTGW instrument). Respondents' scores on the FTGW are reported as a composite for each of the subscales, but there is no composite for the entire instrument. In other words, the subscales are not hypothesized to represent constructs which fall under a superordinate construct, or put another way, there is no hierarchical model for the instrument.

As discussed above in the Introduction, the FTGW was originally developed for use in the English-speaking context, and thus its introduction into the context that is the focus of this study necessitated the translation of the instrument into Japanese. This was done by a process of forward- and back-translation following the International Test Commissions' guidelines (Hambleton, et al., 2005). The translators working on the instrument were all professors or instructors at local universities with some experience in testing. The instrument was first translated into Japanese by a native Japanese speaker in collaboration with the author. The back-translations were undertaken by two native English speakers fluent in Japanese. In the process of comparison between the back-translated version and the original English version, a small number of problematic items were identified. The translations of these items were revised, and all the translators came to agreement on the content of the items. The instrument was considered ready for psychometric evaluation in the Japanese SLA context after the completion of this process.

2.2 Participants and Data Collection Process

A total of 363 respondents took part in this study. The study's participants were students at two universities located in western Japan. Participation in the survey was completely voluntary and formed no part of the respondents' assessment in the class. Informed consent was obtained through the inclusion of a form at the beginning of each questionnaire. This form also clearly stated in Japanese that those students who did not wish to take part in the survey could do so simply by not completing the questionnaire. There was no time limit specified by the administrators of the instrument, however, the time required for the students to complete the questionnaire was approximately 10 minutes in all of the classes to which it was administered In addition to the items comprising the adapted version of the FTGW, respondents were also asked about their group work experience in junior high school, high school and university English classes, and also the frequency of their exposure to group work in these three contexts. The scale for reporting this frequency of exposure included the following seven anchors: never, hardly ever, occasionally, sometimes, often, frequently, or almost always.

From the total 363 records collected from the respondents, 32 were found to have missing data and were therefore removed from the data set. The removal of these records was considered not to have had an effect on the overall properties of the data set as there was no discernable systematic pattern to the missing responses, as determined by inspection by the author. In addition, 24 records were removed from the data set due to an obvious pattern response (e.g., having all 5s checked). The resultant 307 records form the basis for the analysis described in the following section.

Among the respondents, there were 236 males and 70 females (1 respondent did not indicate his/her gender). 53.7% of the respondents were in their 3rd year of study, and the age range of the participants was from 17 to 25. The respondents came from the following fields of study: civil and environmental engineering (n = 62), computer science and electrical engineering (n = 128), applied chemistry and biochemistry (n = 21), medical care and welfare engineering (n = 19), plant science (n = 8), animal science (n = 11), bioscience (n = 7), business administration (n = 20), tourism management (n = 13), and law (n = 18).

2.3 Analytical Procedure

Initially, respondents' scores on the FTGW were entered into a Microsoft Access 2016 database. Calculation of the descriptive statistics and reliability estimates (Cronbach's alphas) for the scores, as well as the EFAs described below, were carried out using IBM/Statistical Package for the Social Science (SPSS) software (Version 21). The primary focus in the initial examination of the data set was from the standpoint of descriptive statistics, with an eye towards the univariate normality (i.e. skew and kurtosis) of scores for items comprising the instrument. These results have been reported previously (see Xethakis, 2018), but are reported here again (Results section below), in abbreviated form, and under citation, because the description of the sample parameters are interpretively important for the subsequent EFAs reported in this section; with the EFA solutions representing new analyses and the central contribution of the paper.

The reliability estimates for the three-factor model originally hypothesized by the authors have also been reported before (see Xethakis, 2018). These are reported here again below for comparative inspection with the new and, as yet, unreported estimates for the three EFA solutions derived in this study.

The EFAs in this study were conducted using the IBM/Statistical Package for the Social Science (SPSS) software (Version 21). The primary aim of EFA is to identify the factors that underlie a set of indicators ― in the case of this study, the indicators are items from a questionnaire ― and to reveal the fewest number of plausible underlying factors needed to explain the relationships (shared variance) found in the data set (Brown, 2015). In this regard, EFA and CFA share a similar overarching objective — to engage with the latent structure that may be found in the data set. However, the two differ greatly in the process of specifying the model which describes this structure As Brown (2015) notes, EFA is exploratory because there are no *a priori* restrictions placed on the pattern of relationships between indicators and underlying factors. In this sense, EFA is a form of post hoc or *a posteriori* analysis: i.e. one that is data-driven and where the number of factors or the pattern of relationships between the underlying factors and the indicators is not initially specified, but rather there to be discovered through a sequence of evidence-based analytic decisions. CFA, on other hand is a form of *a priori* analysis, or direct test, where many of the parameters are specified in advance, chief among these the number of factors and the relationships between indicators and these factors; and this hypothesized solution is evaluated in terms of how well it reproduces, or fits, the relationships in the data set. The model tested most often has either an empirical basis, i.e. from the results of previous studies found in the literature, or a theoretical basis, i.e. a theory-driven conception of which items should indicate which latents and why (typically, but not always, advanced by the authors of an instrument upon first entering the literature). A combination of these two is also possible, and actually desirable.

It is part of the nature of EFA as a method that the results are not determinate, but somewhat contingent upon decisions taken in the process of exploration of the data. For examples, while the results from the initial unrotated solution (Principal Components Analysis) derived for the execution of an EFA are led entirely by the data, and are determinate, the extraction of subsequent and, often rotated solutions, are contingent upon researcher-led determinations of the number of factors to stipulate for extraction. This involves, therefore, researcher judgement and results can be contingent upon this. However, this researcher judgement which should also be evidence-based, and the evidence should come in the form of consideration of the results from the initial solution, usually using procedures such as the eigenvalue-greater-than-one rule and visual examination of a scree plot.

Overall, and importantly, CFA is used when there is an a priori model to test, whereas EFA is typically used under one of two conditions. The first of these conditions is when an instrument is being developed and the dimensionality of the contributive items (i.e. the initial pool of items from which a final subset of the best will be selected) requires empirical exploration. The second of these conditions is when a hypothesized model has failed an a priori test via CFA, and further analysis is required to determine what the dimensional structure of the items comprising the instrument actually is, if not that of the hypothesized model which failed the direct test. The rationale for the EFAs reported in this study comport with the second condition, because Xethakis (2018) reported the failure of the three-factor model, originally hypothesized by Cantwell and Andrews (2002), to plausibly fit the dimensionality of scores produced by the Japanese adaptation of the FTGW in a direct test using CFA as the method.

For the purposes of this study, four separate EFAs were conducted. Three of these were led by the structure of the data as revealed in the process of analysis, and therefore have an empirical rationale, while the fourth EFA was informed by the originally hypothesized three-factor structure of the instrument (Cantwell & Andrews, 2002). This fourth extraction while not having an empirical rational led by the data in this study, was extracted to explore where items would appear if the number of factors was consistent with what the authors hypothesize.

3. Results

The descriptive statistics, previously reported in Xethakis (2018) and re-reported in abbreviated form here, are presented in the first of the three sections below. Following this, the results of the three EFAs are reported. Finally, reliability estimates (Cronbach's alpha) for the factors obtained in the respective solutions for the EFAs are reported for comparative inspection with the previously reported (see Xethakis, 2018) alphas for the three subscales hypothesized by the original authors.

3.1 Descriptive Statistics

The descriptive statistics for the scores summarize the properties of the sample and the distribution of scores for each item. These summaries include the means (as an indicator of central tendency), the standard deviation, and the critical ratio for skew and kurtosis for each item. Table 1 (adapted from Table 1 and Table 2 in Xethakis [2018]) presents a summary of these values. With regard to the means of the items, the highest was that of Item 15 (3.96) while the lowest was that of Item 14 (2.40). The standard deviations of the items varied from 0.739 (Item 13) to 1.097 (Item 02).

Table 1. Item Means, Standard Deviation, Critical Ratios for Skew and Kurtosis for each Item (Absolute Value)

Item	Mean	SD	Skewness Calculated Value	Kurtosis Calculated Value
01	2.63	1.012	1.70	1.58
02	3.41	1.097	**3.61	1.45
03	3.42	0.909	*2.09	0.35
04	2.52	0.818	**4.00	1.22
05	2.79	0.960	1.25	1.50
06	2.88	0.933	0.00	0.92
07	3.44	0.921	**3.01	0.24
08	3.16	0.889	*2.75	0.62
09	2.63	0.858	*2.52	1.51
10	2.67	0.935	1.85	0.83
11	3.37	0.996	1.01	1.03
12	2.62	0.991	*2.48	0.90
13	3.64	0.739	**3.41	1.97
14	2.40	0.839	*2.47	0.17
15	3.96	0.783	**4.83	**3.00
16	2.60	0.885	0.30	0.42
17	2.93	0.969	0.57	1.63
18	2.47	0.930	1.08	1.45

Note: This table is adapted from Table 1 and Table 2 in Xethakis (2018).

With respect to the critical ratios for skew and kurtosis, the resultant values when compared to the stipulated criteria of absolute value 3.0 (relaxed criterion; marked with two asterisks) and absolute value 2.0 (stricter criterion; marked with one asterisk) were problematic in some cases. As can be seen from the results shown in Table 1, in terms of skew, 3 of the 18 items (16.6%) met the threshold to be deemed meritorious (<1.0), and 5 of the items (27.8%) were within the strict threshold (<2.0), while another 5 of the items (27.8%) fell between this value and that of the more relaxed threshold of 3.0. The remaining 5 items (27.8%), however, surpassed the value for the relaxed threshold indicating possible issues with regards to the normality of the data for these items in particular. The items performed better in terms of kurtosis, with 8 of the items (44.4%) meeting the meritorious threshold, 9 items (50.0%) meeting the strict threshold, and the remaining item just on the threshold of the more relaxed value. It must be noted here that in the original study, Cantwell and Andrews (2002) did not report the normality of the scores for the items making up the FTGW and thus it is difficult to determine whether an analogous degree of non-normality was found in the original version of the instrument, or if the source of the non-normality found in the scores on some of the items was introduced in the process of adapting the instrument to the Japanese SLA context.

3.2 Exploratory Factor Analysis

A series of EFAs were conducted in order to explore the possible latent structure of scores in the data set. The rationale for this was that by tentatively identifying factors that underlie the scores on the items of the FTGW, it would be possible to make inferences about what the dimensionality of the scores actually was, given the less than satisfactory outcome of the CFA conducted on the adapted FTGW instrument (Xethakis, 2018). Moreover, it was hoped that the results from the EFA might provide useful guidance for any future revisions of the adapted instrument.

Overall, four EFA solutions were extracted. Three of these were data-led, and conducted on the basis of inferences made from the results an initial PCA, while the fourth was conducted to examine the structure of the data set if the original three-factor model hypothesized by the authors was imposed (Cantwell & Andrews, 2003).

The data set used in this study was initially examined for multivariate normality using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity. The value of the KMO was .828, exceeding the recommended value of .6, while Bartlett's test yielded a significant result (chi^2 [153] = 1271.46, $p < .000$). These two results indicated the multivariate normality of the data and thus its suitability for use in factor analysis.

The first step in the decision-making process for an EFA is to determine the number of factors that should be extracted. This is done by extracting an unrotated solution from the matrix, most often employing PCA. The output of this initial extraction is examined, and a number of means may be used to determine the number of factors to be extracted for the following step of the EFA. The two most commonly employed conventions for making this determination are the eigenvalue-greater-than-one rule and visual inspection of the associated scree plot to ascertain the point of inflection for the slope of the graph.

Table 2 shows the eigenvalues from the initial extraction of factors conducted in this study via PCA. Following the aforementioned eigenvalue-greater-than-one convention, the appropriate number of factors to be specified in the subsequent extraction (rotated) should be five. These factors together explain a cumulative total of 56.27% of the variance found in the data set.

Table 2. Total Variance Explained and Initial Eigenvalues for FTGW

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	4.348	24.157	24.157
2	2.019	11.216	35.373
3	1.447	8.038	43.411
4	1.280	7.114	50.524
5	1.034	5.744	56.268
6	0.940	5.223	61.491
7	0.838	4.655	66.147
8	0.802	4.455	70.602
9	0.677	3.760	74.362
10	0.649	3.606	77.968
11	0.630	3.502	81.470
12	0.591	3.283	84.753
13	0.573	3.181	87.933
14	0.541	3.006	90.939
15	0.483	2.681	93.620
16	0.447	2.483	96.104
17	0.364	2.020	98.124
18	0.338	1.876	100.00

Extraction Method: Principal Components Analysis

Figure 1 shows the scree plot generated by this initial extraction of factors for the FTGW. The slope of the plot begins to flatten out after the fifth component, which again suggests that a five-factor solution would be appropriate. This result aligns with the eigenvalue greater than one convention, and therefore a five-factor solution was initially attempted.

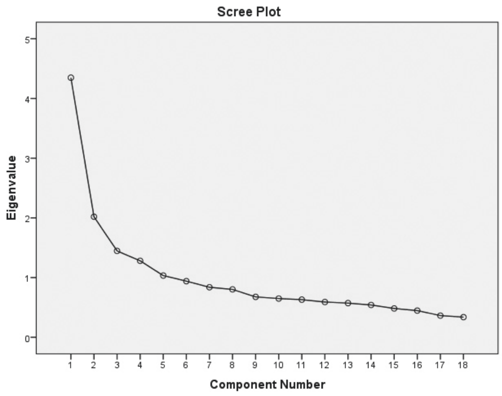

Figure 1. Scree plot of the eigenvalues generated by unrotated extraction (PCA) of factors from the data matrix of FTGW scores.
Note: The vertical axis shows the eigenvalues, while the horizontal axis indicates the component number (1 - 18, one for each item on the FTGW). The small circles show the eigenvalue of each component, with the line connecting the circles representing the slope (extent) of the difference between component values.

The results of the initial unrotated extraction employing PCA are presented in Table 3. Component 1 loads the 7 items of the PIL subscale most strongly, and tends to load the 7 items of the PGL negatively. This result appears to be in line with the higher alpha value for this subscale reported in Xethakis (2018). Interestingly, two of the items of the DGL subscale also load strongly on this component. The majority of the items from the PGL (4 of 7) load onto Component 2, with a few negative loadings for items on the PIL scale. Component 3 loads two of the items for the PGL strongly, but otherwise appears to weakly load items in a more random manner. The fourth component loads the remaining two items of the DGL subscale, but again the other items load weakly and randomly. The fifth and final component loads one item of the PGL subscale quite strongly and one item from the DGL weakly, but is otherwise random in its array of positive and negative loadings. This distribution of items among the extracted components follows the methodology of PCA, where a factor is extracted which accounts for the largest share of variance, and then from the residual variance a subsequent factor is extracted which accounts for the largest share of the remaining variance, and so on. This is why the first factor is so dominant in this unrotated solution.

Table 3. Component Matrix for FTGW

	Component				
	1	2	3	4	5
Item 01 (PIL)	0.704	-0.160	0.318	-0.110	0.006
Item 02 (DGL)	0.258	0.192	0.315	0.678	-0.056
Item 03 (PGL)	-0.445	0.324	0.028	0.320	-0.224
Item 04 (DGL)	0.341	-0.083	-0.172	0.515	0.421
Item 05 (PIL)	0.667	0.242	0.180	-0.187	-0.018
Item 06 (PIL)	0.617	-0.012	0.295	-0.063	-0.126
Item 07 (PGL)	-0.513	0.463	-0.044	0.028	0.143
Item 08 (PGL)	-0.086	0.670	-0.093	0.144	-0.044
Item 09 (PGL)	0.054	0.649	-0.425	-0.204	-0.077
Item 10 (DGL)	0.454	0.367	0.089	0.109	-0.243
Item 11 (PGL)	0.073	0.359	0.250	-0.126	0.736
Item 12 (PIL)	0.815	0.039	0.099	-0.052	-0.068
Item 13 (PGL)	-0.254	0.342	0.448	-0.448	-0.173
Item 14 (PIL)	0.559	-0.021	-0.309	-0.183	0.207
Item 15 (PGL)	-0.402	0.331	0.533	-0.009	0.189
Item 16 (PIL)	0.636	0.163	-0.201	-0.15	0.131
Item 17 (DGL)	0.614	0.121	0.191	0.223	-0.135
Item 18 (PIL)	0.420	0.410	-0.410	0.038	-0.099

Extraction Method: Principal Component Analysis 5 components extracted

While the results for this initial unrotated extraction are reported here, they should not serve as the basis for any final analysis of the latent structure of the data, due to the nature of the unrotated PCA used to derive them with its dominant first factor, and noted above. The final interpretation of the data should be based on the rotated solution and its associated pattern matrix (if the form of the rotation is oblique), which are described below for each of the EFAs conducted in this study. There are five factors in the unrotated solution shown in Table 3, because the eigen-value-greater-than-one-rule (the default option for choosing the number of factors in SPSS) was used for this unrotated solution.

Following the determination of the number of factors to be extracted on the basis of the eigenvalue-greater-than-one rule and the scree plot, a five-factor extraction was initially conducted on scores for the FTGW using Direct Oblimin rotation. This is a form of oblique rotation that allows the factors to correlate. Correlation of factors is a common supposition in the social sciences, where factors are assumed not to be truly orthogonal, or uncorrelated, as they might be with data from the natural sciences (see Kline, 1994).

Table 4 presents the results of this initial five-factor extraction. The loading threshold for this extraction was set at greater than .40, and thus Table 6 shows only those items which loaded on their respective factors with a coefficient value greater than absolute value .4.

Table 4. Pattern Matrix for Five-factor Extraction on FTGW

	Factor				
	1	2	3	4	5
Item 05 (PIL)	0.964				
Item 06 (PIL)	0.580				
Item 01 (PIL)		-0.654			
Item 07 (PGL)		0.642			
Item 12 (PIL)		-0.476			
Item 03 (PGL)		0.419			
Item 09 (PGL)			0.690		
Item 18 (PIL)			0.536		
Item 02 (DGL)				0.666	
Item 15 (PGL)					0.686
Item 13 (PGL)					0.467
Item 04 (DGL)					
Item 08 (PGL)					
Item 10 (DGL)					
Item 11 (PGL)					
Item 14 (PIL)					
Item 16 (PIL)					
Item 17 (DGL)					

Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization
Rotation converged in 21 iterations.

While the data from the five-factor result is presented in this paper, it must be noted that this extraction produced an instance of an improper solution. More specifically this solution generated a Heywood case, or a solution with a parameter estimate that contains an implausible value (Kline, 2011). In the case of this attempted five-factor solution, the out-of-range value derived was a communality value greater than 1. Communality can be characterized as the proportion of the variance found in a particular item that can be explained by the underlying factor (Brown, 2015). Clearly, the amount of variation in an item explained by the latent variable cannot be greater than one, and thus the results generated by this five-factor solution should, at least, be interpreted with a great degree of caution, if not rejected outright. In this study, the latter course was adopted, and for this reason, both a six-factor and a four-factor EFA were conducted following the rejection of the five-factor solution.

The resulting six-factor extraction again yielded an improper solution with a Heywood case similar to that of the five-factor extraction and so was rejected as well (the results for this solution are not presented). Table 5 presents the results of the rotated four-factor extraction (Direct Oblimin), again with the threshold for loading set at greater than absolute value .4.

Table 5. Pattern Matrix for Four-factor Extraction on FTGW

	Factor			
	1	2	3	4
Item 01 (PIL)	0.755			
Item 12 (PIL)	0.745			
Item 05 (PIL)	0.688			
Item 06 (PIL)	0.654			
Item 17 (DGL)	0.442			
Item 16 (PIL)	0.441			
Item 03 (PGL)	-0.427			
Item 07 (PGL)	-0.423			
Item 09 (PGL)		0.699		
Item 18 (PIL)		0.511		
Item 08 (PGL)		0.429		
Item 02 (DGL)			0.702	
Item 13 (PGL)				0.573
Item 15 (PGL)				0.516
Item 04 (DGL)				
Item 10 (DGL)				
Item 11 (DGL)				
Item 14 (PIL)				

Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization
Rotation converged in 17 iterations.

The results of this solution show that 5 of the 7 items of the PIL subscale load on Factor 1, several quite strongly, while two of the items from the PGL subscale also load negatively on this factor. Interestingly, one item from the DGL subscale also loads on this factor. Factor 2 consists of 2 items from the PIL subscale and one from the PGL subscale. Only one item loads onto Factor 3, an item from the DGL subscale. Two items from the PGL subscale load onto Factor 4. Finally, four items from the FTGW failed to load on any factors — two from the DGL subscale and one each from the PIL and DGL subscales.

In addition to the three data-driven EFAs conducted — the five-, six- and four-factor solutions above — a three-factor solution was also extracted. The choice of the number of factors in this solution was informed by the three-factor solution originally proposed by Cantwell and Andrews (2002) on the basis of their own EFA. Table 6 shows the results of this rotated extraction employing Direct Oblimin, again with a loading threshold of absolute value .4.

Table 6. Pattern Matrix for Three-factor Extraction on FTGW

	Factor		
	1	2	3
Item 01 (PIL)	0.733		
Item 12 (PIL)	0.728		
Item 05 (PIL)	0.663		
Item 06 (PIL)	0.638		
Item 17 (DGL)	0.535		
Item 09 (PGL)		0.668	
Item 18 (PIL)		0.492	
Item 08 (PGL)		0.46	
Item 15 (PGL)			0.63
Item 13 (PGL)			0.469
Item 02 (DGL)			
Item 03 (PGL)			
Item 04 (DGL)			
Item 07 (PGL)			
Item 10 (DGL)			
Item 11 (PGL)			
Item 14 (PIL)			
Item 16 (PIL)			

Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization
Rotation converged in 10 iterations.

In this three-factor solution, four of the seven items from the PIL subscale and one item from the DGL subscale loaded on Factor 1. This is quite similar to the result from the four-factor solution where the same four items of the PIL subscale and the same item from the DGL subscale loaded onto the first factor extracted in that solution. Factor 2 included two items from the PGL subscale and one from the PIL subscale. This result is again quite similar to that of the four-factor solution, where these three items clustered together on the second factor extracted. The third factor contains two items from the PGL subscale. These two items loaded onto the forth factor extracted in the four-factor solution. The eight remaining items failed to load (i.e. with coefficients above absolute value .4) in this solution. Four of these, Items 4, 10, 11 and 14, failed to load on the four-factor solution as well. Item 16, which loaded on the first factor in the four-factor solution, did not load on any factor in this three-factor version. Two items, Items 3 and 7, loaded negatively on the first factor extracted in the four-factor solution but failed to load in this solution. Finally, Item 2, which comprised the third factor in the four-factor solution, failed to load on any factor in this three-factor solution.

3.3 Reliability Estimates

Table 7 presents the reliability estimates (Cronbach's alpha) for scores on each of the factors extracted in the four- and three-factor solutions reported above. The highest alpha value was that for Factor 1 (.81) of the four-factor solution. Factor 3 of the four-factor solution comprised only a single item and so no value is reported. The values for Cronbach's alpha for each of the three subscales were reported in the original study as .78 (PIL), .71 (PGL) and .60 (DGL), respectively.

Table 7. Reliability Estimates for Scores on Extracted Factors for Four- and Three-Factor Solutions

Solution	Cronbach's alpha Factor Number			
	1	2	3	4
Four-factor	.81	.52	−	.46
Three-factor	.80	.52	.46	−

As can be seen in Table 7, the reliability estimates of the first factor in both the four- and three-factor solutions (which as noted above are comprised of almost the same items) exceed the benchmark (.70) recommended by Nunnally and Bernstein's (1994), and thus the factors would appear to meet the minimum standard to be considered a reliable measure on the basis of this index. The possibility that the relatively high values for alpha on these two factors results from the content of the items that loaded on these factors is examined in the Discussion section below.

The remaining factors in the four- and three-factor solutions did not reach the .70 threshold and these results could bring the reliability of scores on these factors into question. However, as Cortina (1993) and Green, Lissitz and Muliak (1977) have shown, alpha tends to reward subscales with a larger number of items. Factor 2 and Factor 4 of the four-factor solution have three and four items respectively and so the small number of items could be one of the reasons for the low alpha value. Factor 2 and Factor 3 of the three-factor solution, are comprised of the same items as the second and fourth factor of the four-factor solution, and thus the same arguments apply to these two factors as well.

4. Discussion

The purpose of the EFA described in this paper was twofold. First, use of the methodology aimed to explore the dimensionality of the scores in the data set, given the less than satisfactory results obtained from a CFA (Xethakis, 2018) undertaken to ascertain the degree of fit between the scores on a version of the FTGW adapted for the Japanese EFL context and the structure of the instrument originally hypothesized by Cantwell and Andrews (2002). Second, and through this exploration, the aim of the EFA was to aid in any possible future development of the FTGW for use in the Japanese population. This would include the possible development of a new scale aimed at identifying learners who might be 'at risk' in group-based learning approaches due to feelings of anxiety or other forms of psychological or cognitive distress. The DGL subscale, which exhibited meritorious fit in the diagnostic CFA reported by Xethakis (2018), might have the potential to serve as the core of such a scale.

The initial unrotated extraction suggested a five-factor solution, based on the eigen-value-greater-than-one rule (See Table 2) as well as the scree plot (See Figure 1). This is greater than the number of factors, i.e. three, in the originally hypothesized model (Cantwell & Andrews, 2002). Based on this evidence a five-factor solution was extracted using Direct Oblimin rotation. This solution, however, was abandoned on the basis of two issues that arose in the course of the extraction. First, as mentioned above, the extraction produced a communality value greater than 1, which brought the solution into question, and moreover the solution failed to converge within the usual 25 iterations. A further solution was attempted with a greater number of iterations permitted (50) and the solution converged in 29 iterations, however the communality problem remained, which created issues for the interpretation of the results of this extraction.

These issues led to a re-examination of the scree-plot, and on the basis of this re-inspection, it was decided to attempt two further extractions ― a six- and a four-factor extraction. The six-factor extraction converged rather quickly, in 16 iterations. However, this solution also generated a Heywood case, again a case of a communality value greater than 1, and so was abandoned as problematic for interpretation. The four-factor solution led to an extraction which converged after 17 iterations and did not engender any Heywood cases.

The patterns of item loadings on the individual factors for the four-factor solution, with the criterion for an acceptable loading being a coefficient value of greater than absolute value .4, differed noticeably from those found by Cantwell and Andrews (2002). Factors 1 and 2 contained items from at least two of the originally hypothesized subscales (See Table 5). The third factor was comprised of only one item from the DGL subscale while the remaining factor was made up of two items from the PGL subscale. Factor 1 had a total of eight items. While these items were from all three of the subscales, the majority of these (5 of 8 the items on this factor) came from the PIL subscale. It is interesting to note that the two items from the PGL subscale which grouped on this factor did so negatively. This would be expected under the presumption that preference for group and individual work are opposite ends of a single bi-polar construct, rather than entirely distinguishable constructs. The second factor extracted consisted of two items from the PGL subscales and one item from the PIL subscale, which loaded positively rather than negatively; a negative loading would be expected, under the same presumption as mentioned above, and indicating a negative orientation to the other two PGL items which should be dominant given that there are two of them. Factor 3 was a single-item factor, with that item coming from the DGL subscale. The forth factor had two items from the original PGL subscale. It should also be noted that four items, two from the DGL subscale, and one each from the PIL and PGL subscales, failed to load above the absolute value .4 threshold.

In order to explore possible reasons for the lack of structural correspondence between the original model (Cantwell & Andrews, 2002) and the model suggested by this four-factor solution, the content of the items under the four-factor solution was inspected more closely.

The first factor, with five of its eight items coming from the PIL subscales, seems to preserve the core of the subscale. Additionally, the relatively high alpha value for this factor (.81) was roughly commensurate with previous values obtained for the PIL subscale (but with only the originally hypothesized seven items representing it); e.g. .78 in the original study (Cantwell & Andrews, 2002), .79 in Xethakis (2018), and .82 in White, et al., (2005). The alpha value for this particular factor was also the highest of those to emerge from factors produced by the EFA solutions pursued in this study.

It is arguable that the content of the items on the

PIL subscale contribute substantially to their co-occurrence on the factor, and to the associated its high alpha value. Items 1, 5, 6 and 12 all express a dislike for group work, as is to be expected from items which were originally ascribed by Cantwell and Andrews to be part of a subscale expressing a preference for individual learning. However, it should be noted that these four items express a rather similar sentiment. In fact, Items 1 and 12 could be seen as opposites, as could Items 5 and 6, which is why one of each pair require reverse scoring on the original FTGW (Cantwell & Andrews, 2002). After these two items are reversed scored, there are four items having very similar content and this may be a large part of why the original PIL subscale consistently exhibits the relatively high reliability estimates that have been reported. This is also a possible reason why these estimates may be considered as artificially inflated because of the repetitive nature of their content. This similarity in item content also raises the question of whether all four items are really needed in the PIL subscales and whether the two reverse scored items might be removed to create a more parsimonious subscale.

Returning to the other items on the first factor, Items 7 and 17 are from two different subscales, the PGL and DGL, respectively. However, both appear to be related to affective aspects of group work. Item 17 expresses a general feeling of anxiety when working in groups, as would be expected with an item originally from the DGL subscale, and is therefore negatively oriented to group work, and thus conversely, positively oriented to individual work and hence its appearance in this factor. Item 7, however, expresses a sense of satisfaction with group accomplishments, and it would therefore be difficult to account for its appearance on this factor given the content of other items with which it is somewhat incommensurate, were it not for the fact that it is negatively oriented to the factor. This makes its place on the factor quite commensurate with the content of the other items. Its negative loading on this factor expresses fundamentally an aversion to group work, or perhaps apprehensiveness about working in groups. Item 3, which loaded negatively, and Item 16, are both concerned with cognitive demands inherent in group work, but are from separate subscales, the PGL and PIL, respectively. Nonetheless, both express a sense of confusion or lack of understanding that may come with working in groups.

From an examination of item content, it seems that this factor consists of items that express a general dislike of group-based learning as well as some of the cognitive complexities and affective elements that come with it. For this reason, it may be possible to interpret this particular factor as expressing an aversion to group work, rather than a preference for individual work, and again this centers the analytical and empirical burden of determining whether preferences for group and individual work represent separate constructs, or opposite poles of a single construct. The preliminary evidence in this paper would seem to suggest the latter.

The second factor extracted from the four-factor solution seems to express a positive evaluation of several aspects of group work such as individual effort on behalf of the group, positive interdependence and collective responsibility. It may be possible to interpret this three-item factor as relating to a sense of commitment to group work.

Factor 3 consists of just one item, Item 2, originally hypothesized as an indicator of the DGL subscale. The item expresses unease with two aspects of group work, expressing opinions and communicating with others. However, it is not clear that these two aspects reflect a single overarching concern. The cause of learners' unease might be simply speaking in the front of others in the group which while stressful would seem to be an indicator of social anxiety, and may not give rise to further fears of negative evaluation that could come with having to express one's opinions in a group. These two aspects might, in fact, represent two discrete latent constructs. If this were the case, it would suggest that this particular item could be revised into two separate items with each expressing one of these constructs. In principle, an item should not be measuring more than one latent construct. Moreover, in the Japanese context, it may be that respondents are more sensitive to expressing opinions in general than specifically in the group context.

The fourth factor while consisting of two items, Item 13 and 15, from the PGL subscale, seems to express two separate aspects of group work. Item 13 deals with cognitive demands, while at the same time, Item 15 seems concerned more with aspects of positive interdependence. The common underlying construct seems rather unclear in this case.

In light of the differences found between the grouping of the items in this evidence-based four-factor extraction and those of the original FTGW, an additional rotated solution with three-factors was extracted to investigate the possibility that the items might load in a more coherent fashion when restricted to three factors. This extraction produced a solution that was in many respects quite similar to the four-factor solution (See Table 6). Items which grouped together on Factor 2 in the four-factor solution also grouped together on the second factor in the three-factor solution and at quite similar strengths (i.e. relatively similar coefficients). The two items from the fourth factor remained together on Factor 3 in the three-factor solution, again at comparable strengths. The primary differences between the two extractions were the elimination of the single-item factor (Factor 3 in the four-factor solution) and the loss of three items from the first factor extracted, and a consequent increase in the number of items (8) that failed to load on any factor at a threshold of absolute value .40.

It is interesting to note that the two items from the PGL subscale (Items 3 and 7) which loaded negatively onto the first factor in the four-factor solution failed to load onto Factor 1 in this three-factor solution, as did one item from the PIL subscale (Item 16). The loss of these three items seems to substantially narrow the operational expression of the construct which would be presumed to underlie this factor. Aversions to cognitive complexity and the lack of a sense of satisfaction in group accomplishment are absent in the factor. What remains in this factor seems to be a general dislike of working with others and a vague sense of anxiety at being placed in a group.

Overall, it seems that the three-factor extraction, rather than revealing a more coherent structure, in the sense of resembling Cantwell and Andrews (2002) original conception, confirms aspects of the four-factor solution, while at the same time, reduces the range of operational expression of some of its presumed latent constructs.

From the above, it would seem that the two plausible, and indeed possible, EFA solutions extracted in this study were less than satisfactory in presenting a clear picture of the dimensionality of the scores gathered. A large number of items failed to load on any factors and moreover, the number of items loading at an absolute value of .4 on most of the factors was quite small ― three or fewer items on five of the seven factors across the two solutions. This is less even than on the shortest subscale in the original FTGW (4 items) obtained by Cantwell and Andrews (2002) which was the DGL. Finally, the amount of variance explained by the two solutions is less than might have been hoped. The components extracted in the four-factor solution explained 51% of the variance in the data set, while those of the three-factor solution explained 43%, leaving 49% and 57%, respectively, of the variance unexplained. This unexplained variance will inevitably include error (i.e. noise) in the scores comprising the data set, of which there will always be some. However, and importantly, it may also include systematic variance from latent variables that were either 1) not specified in the originally hypothesized structure of the FTGW (Cantwell & Andrews, 2002) or 2) which have emerged in the process of adaptation into the Japanese EFL context.

One possible reason for the lack of similarity between the original FTGW and the adapted version examined in this study, and which goes beyond the translation into Japanese and re-specification of the domain as EFL, may be the move from secondary- to tertiary-level populations. Perhaps the coherence displayed in the subscales from Cantwell and Andrews' (2002) study reflected to a greater degree the concerns and dispositions of secondary students. These would include issues such as a greater fear of expressing oneself in front of others, the possibility of negative evaluations by peers, as well as the greater degree of cognitive complexity experienced by adolescents in group work (Cantwell, 1998).

Some evidence for this view might be found in the literature concerning the FTGW. In addition to the less than satisfactory outcome of the CFA in Xethakis (2018), one other CFA has been performed, that of Goudas, et al. (2009). As mentioned in the Introduction section above, the results of the CFA showed a good degree of fit between Cantwell and Andrews' (2002) hypothesized structure and scores on the G-FTGW. However, it must be noted that the population of interest in this study was students in the sixth-grade. Moreover, the instrument used in this study was substantially modified from Cantwell and Andrews' (2002) instrument. It could be argued that the degree of modification observed in the G-FTGW means that the good fit obtained for this version of the instrument does not speak to the level of fit one could expect from the original version were it to be tested via CFA in the population for which it was proposed. While EFA can also provide some evidence for the coherence of the subscales on an instrument, there have only been two EFAs conducted on the FTGW: that of Cantwell and Andrews' (2002) pilot study and that of this study, which does not confirm the findings of the original study. The case could be made, of course, that the lack of coherence found in the results of this study is an artifact of the adaptation process. However, an argument could also be made that this lack of correspondence relates to the actual subscales of the original instrument when used in university settings rather than in secondary school settings. Adjudicating between these two cases for explanation of misfit would necessitate further research.

Another piece of negative evidence for the coherence of the subscales on the FTGW might be found in the low reliability estimates for these subscales when employed in university settings. In the four studies which have reported reliability estimates (Forrester & Tashchian, 2010; Gasaymeh, et al., 2014; White, et al., 2005; and Xethakis, 2018), the alpha values for the PGL subscales has exceeded Nunnally and Bernstein's (1994) .70 benchmark only once — in Gasaymeh, et al., (2014) where the value for the PGL subscale was .73, just above the threshold. Even in the original study the value for this scale was only .71. The DGL subscale has fared even worse, producing values of between .66 (Gasaymeh, et al., 2014) and .52 (Xethakis, 2018). This subscale also performed poorly in the original study with a value of only .60. It should be noted that in Cantwell and Andrews' (2002) study the reliability estimate for the PIL subscale performed better than the other two subscales, with a reported value of .78. As mentioned above, the alpha value for this subscale was also relatively high (.82) in the one study at the tertiary level which reported values for this subscale (White, et al., 2005). However, the instrument used in White, et al. was an adapted version of the FTGW, with items removed specifically with the intent of raising the reliability estimates of the subscales (White, et al., 2005, p. 619).

A final point which should be touched on in this discussion is the performance of the DGL subscale, which failed to show a strong degree of coherence in this study when compared with its meritorious performance in the diagnostic CFA conducted by Xethakis (2018). In this study, only one item from the original DGL subscale (Item 17) loaded at an absolute value of .40 in both of the two solutions presented here. An additional item from the original subscale (Item 2) loaded as a solitary item on Factor 3 in the four-factor solution, while two of the items from the subscale (Items 4 and 10) failed to load on any factor in either solution. One possible reason for the pattern of loadings outlined above may be the comparative sizes of the three subscales. The PIL and PGL each comprise seven items as opposed to the DGL's four. It should be noted that the dimensionality of the DGL subscale was confirmed when examined as a single scale under CFA, but was not readily apparent when the dimensionality of the scores were tested alongside companion subscales on the FTGW as a whole.

5. Conclusion

From the above, it would seem that the aim of clarifying the dimensionality of the scores gathered in this data set, points to a lack of coherence in general, rather than a coherence different from Cantwell and Andrews (2002). However, the results of the EFA do also seem to bear on critical theoretical issues with respect to conceptualizing the measurement model for the instrument, as well as point the way to a number of means by which this adapted version of the FTGW could be adjusted in any future research.

One general issue concerns the question of whether preference for group work and preference for individual work should be treated as two separate constructs, or as opposite poles on a single construct, as discussed above.

The results also arguably appear to show two possible dimensions underlying the scores in this data set. One of these dimensions was an affinity towards group commitment. This aspect of group work made up a part of the PGL subscale in Cantwell and Andrews' (2002) study, but perhaps in the Japanese population this may represent an independent dimension, or is the only part of the original construct finding expression in this population.

The second of the dimensions revealed in this study was an aversion to group work, with its attendant cognitive complexity and affective demands. This corresponds with findings (e.g., Cantwell & Andrews, 2002; Forrester & Tashchian, 2010) showing that students who were less able to deal with the cognitive or affective demands found in group-based learning approaches also tended to be averse to group work. This suggests that one avenue for further research could be a deeper examination of the particular cognitive and affective factors that underlie learners' aversions to group work.

In addition, the prominence of this dimension in the present study, taken in combination with the findings of Xethakis (2018), suggests the possibility that this dimension and the DGL subscale, rather than representing completely distinct constructs, represent a continuum of negative feelings toward group work, at least in this population.

Finally, in regards to the limitations of this study as a whole, it must be noted that this study deals with just one sample. Moreover, the sample is not a truly representative sample of the target population ― Japanese EFL students ― but a sample of convenience, which limits the generalizability of the results. Further studies sampling from the same population could help to ameliorate this limitation by giving broader representation to the population of interest.

References

Angeli, C., & Tsaggari, A. (2016). Examining the effects of learning in dyads with computer-based multimedia on third-grade students' performance in history. *Computers & Education, 92*, 171-180.

Archer, J. (1994). Achievement goals as a measure of motivation in university students. *Contemporary Educational Psychology, 19*(4), 430-446.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Cantwell, R. H., & Andrews, B. (2002). Cognitive and psychological factors underlying secondary school students' feelings towards group work. *Educational Psychology, 22*(1), 75-91.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98-104.

Costa, P. T. and McCrae, R. R. (1992). *Neo PI-R and Neo-FFI Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Dörnyei, Z., & Malderez, A. (1997). Group dynamics and foreign language teaching. *System, 25*(1), 65-81.

Fantuzzo, J. W., Riggio, R. E., Connelly, S., & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. *Journal of Educational Psychology, 81*(2), 173-177.

Forrester, W. R., & Tashchian, A. (2010). Effects of personality on attitudes toward academic group work. *American Journal of Business Education, 3*(3), 39.

Fushino, K. (2010). Causal relationships between communication confidence, beliefs about group work, and willingness to communicate in foreign language group work. *TESOL Quarterly, 44*(4), 700-724.

Gasaymeh, A. M. M., Kreishan, L. J., & Al-Dhaimat, Y. A. (2014). The effects of age, gender, and major on measures of Jordanian students' attitudes towards academic group learning in computer training courses. *Journal of Education and Practice, 5*(15), 173-179.

Gillies, R. M. (2003). Structuring cooperative group work in classrooms. *International Journal of Educational Research, 39*(1), 35-49.

Gorvine, B. J., & Smith, H. D. (2015). Predicting student success in a psychological statistics course emphasizing collaborative learning. *Teaching of Psychology, 42*(1), 56-59.

Goudas, M., & Magotsiou, E. (2009). The effects of a cooperative physical education program on students' social skills. *Journal of Applied Sport Psychology, 21*(3), 356-364.

Goudas, M., Magotsiou, E., & Hatzigeorgiadis, A. (2009). Psychometric properties of the Greek version of the Feelings Toward Group Work questionnaire. *European Journal of Psychological Assessment, 25*(3), 204-210.

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37*(4), 827-838.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Erlbaum.

Jackson, D. (1974). *Personality Research Form Manual*. Goshen, NY: McGraw Hill.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college what evidence is there that it works? *Change*: *The Magazine of Higher Learning, 30*(4), 26-35.

Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). NY: Guilford Press.

Kondo, D. S., & Yang, Y .L. (1994). Strategies for reducing social anxiety. *Communication Research Reports, 11*(2), 153-159.

Lemonia, D., & Dimitris, G. (2017). Cooperative learning as a preference of adolescent students: the case of dance activities. *Academic Journal of Interdisciplinary Studies, 5*(3 S1), 478.

Maesin, A., Mansor, M., Shafie, L. A., & Nayan, S. (2009). A study of collaborative learning among Malaysian undergraduates. *Asian Social Science, 5*(7), 70.

Ministry of Education, Culture, Sports, Science and Technology. (2008). Section 9 Foreign Languages. In *Improvement of Academic Abilities (Courses of Study)*. Retrieved July, 9th, 2018 from http://www.mext.go.jp/component/english/__icsFiles/afieldfile/2011/03/17/1303755_013.pdf

Ministry of Education, Culture, Sports, Science and Technology. (2015). *National University Reform for the Coming Era* Retrieved July, 9th, 2018 from http://www.mext.go.jp/en/policy/education/highered/title02/detail02/__icsFiles/afieldfile/2015/10/01/1362381_1_1.pdf

野崎秀正 [Nosaki Hidemasa]. (2012). 協同学習が仲間への質問行動に対する態度に及ぼす影響 [The Effects of Cooperative Learning on Students' Attitudes toward Asking Questions to Peers]. 宮崎学園短期大学紀要 [*Bulletin of Miyazaki Gakuen Junior College*], *5*, 121-130.

野﨑秀正 [Nosaki Hidemasa]. (2016). 授業中の協同活動に対する認識の学年差と達成目標の関連　[The Effects of Achievement Goals on Perceived Cooperation in Elementary and Junior High School Students]. 宮崎公立大学人文学部紀要 [*Bulletin of Miyazaki Municipal University Faculty of Humanities*], *23*(1), 53-60.

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory, 3*, 248-292.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460-475.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21(1), 43-69.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*(4), 448-457.

White, F., Lloyd, H., Kennedy, G., & Stewart, C. (2005). An investigation of undergraduate students' feelings and attitudes towards group work and group assessment. *Research and Development in Higher Education, 28*, 618-623.

Xethakis, L. (2018). Psychometric adaptation of a Japanese version of the Feelings Towards Group Work questionnaire for use in the Japanese SLA context. *Kumamoto University Studies in Social and Cultural Science, 16*, 219-247.

Zhou, M. (2016). The roles of social anxiety, autonomy, and learning orientation in second language learning: A structural equation modeling analysis. *System, 63*, 89-100.

Appendix — Feelings Towards Group Work Instrument (Cantwell & Andrews, 2002)

This questionnaire examines your preferences in relation to working and studying in groups.

Please read each statement and then indicate by circling the appropriate number, the degree to which you think the statement is true of you. If you believe the statement to be very true of you, circle the "5". If you believe the statement to be not true of you at all, then circle the "1". If you believe you are somewhere between these extremes, circle the "2", "3" or "4".

(1) I enjoy working within a group.
(2) I sometimes feel nervous when I have to give my ideas or communicate within a group.
(3) I understand information better after explaining it to others in a group.
(4) I often find it difficult to understand what the group task is.
(5) I like to work alone even when placed in a group.
(6) I prefer to work within a group rather than work alone
(7) I often have a strong feeling satisfaction when I become totally involved in a group achievement.
(8) It is important that other group members take responsibility for my learning as well.
(9) I usually make a strong personal contribution to group work.
(10) I am often afraid to ask for help within my group.
(11) I like group work more when we can make up our own groups.
(12) I do not like to study within a group.
(13) I can usually understand other group members' ideas.
(14) Even when groups are well organised, I don't believe they are a more effective way of using class time.
(15) It is best when each person helps each other within a group.
(16) I often think the work becomes too confusing when done in a group rather than individually.
(17) I rarely feel relaxed within a group.
(18) I sometimes feel let down by other group members.

Preference for Individual Learning subscale (PIL): Items 1, 5, 6, 12, 14, 16 and 18

Preference for Group Learning subscale (PGL): Items 3, 7, 8, 9, 11, 13 and 15

Discomfort in Group Learning subscale (DGL): Items 2, 4, 10 and 17